

300

| | | |
|------------------|-----------------|----------------------------------|
| Cookies ~ 302 | Tokens ~ 304 | Occurrence Frequency ~ 306 |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

Figure 3a

310

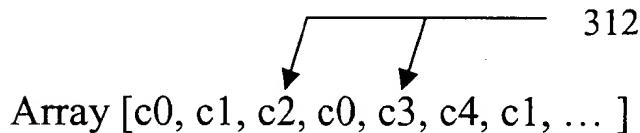


Figure 3b

320

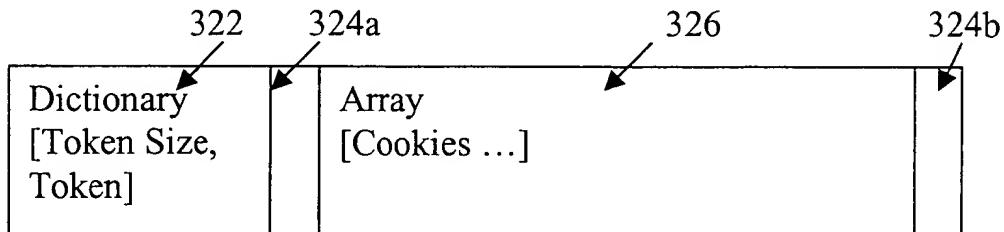


Figure 3c

430a

Array[] = { 0, 1, 2, 0, 3, 5, 6, ... }

Figure 4d

420b

| <u>Cookie</u> | <u>Tokens</u> | <u>Frequency</u> |
|---------------|---------------------|------------------|
| 1 | ">" | (8) |
| 2 | "=" | (6) |
| 3 | "Employee" | (6) |
| 4 | "<" | (4) |
| 5 | "</" | (4) |
| 6 | "ID" | (3) |
| 7 | "Software Engineer" | (3) |
| 8 | "Title" | (3) |
| 9 | "Employees" | (2) |
| 10 | "1" | (1) |

...

Figure 4e

430b

Array[] = { 4, 8, 1, 4, 3, 6, 10, 8, 7, ... }

Figure 4f

440

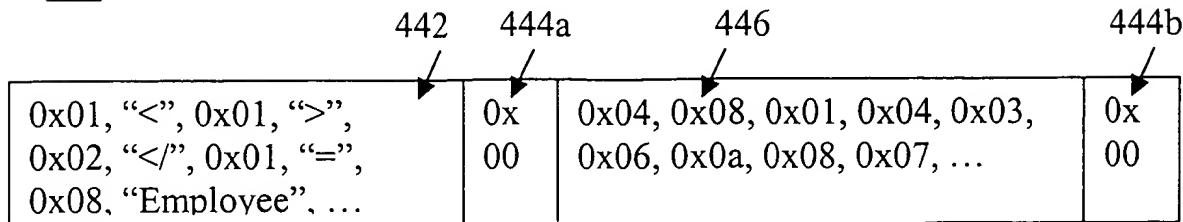


Figure 4g